

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/522,789  
Source: PG/10  
Date Processed by STIC: 2/8/05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,789

DATE: 02/08/2005

TIME: 11:42:55

Input Set : N:\Rasheed\PTO.SR.txt

Output Set: N:\CRF4\02082005\J522789.raw

```

3 <110> APPLICANT: TANABE SEIYAKU CO., LTD.
5 <120> TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
7 <130> FILE REFERENCE: 03-039-PCT
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/522,789
C--> 9 <141> CURRENT FILING DATE: 2005-01-28
9 <150> PRIOR APPLICATION NUMBER: US 60/398,761
10 <151> PRIOR FILING DATE: 2002-07-29
12 <160> NUMBER OF SEQ ID NOS: 2
14 <170> SOFTWARE: PatentIn version 3.1
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 2301
18 <212> TYPE: DNA
19 <213> ORGANISM: Homo sapiens
21 <220> FEATURE:
22 <221> NAME/KEY: CDS
23 <222> LOCATION: (1)..(2301)
24 <223> OTHER INFORMATION:
W--> 26 <400> 1
27 atg aag aca ccg tgg aag gtt ctt ctg gga ctg ctg ggt gct gct gcg      48
28 Met Lys Thr Pro Trp Lys Val Leu Leu Gly Leu Leu Gly Ala Ala Ala
29 1          5          10          15
30 ctt gtc acc atc atc acc gtg ccc gtg gtt ctg ctg aac aaa ggc aca      96
31 Leu Val Thr Ile Ile Thr Val Pro Val Val Leu Leu Asn Lys Gly Thr
32          20          25          30
33 gat gat gct aca gct gac agt cgc aaa act tac act cta act gat tac      144
34 Asp Asp Ala Thr Ala Asp Ser Arg Lys Thr Tyr Thr Leu Thr Asp Tyr
35          35          40          45
36 tta aaa aat act tat aga ctg aag tta tac tcc tta aga tgg att tca      192
37 Leu Lys Asn Thr Tyr Arg Leu Lys Leu Tyr Ser Leu Arg Trp Ile Ser
38          50          55          60
39 gat cat gaa tat ctc tac aaa caa gaa aat aat atc ttg gta ttc aat      240
40 Asp His Glu Tyr Leu Tyr Lys Gln Glu Asn Asn Ile Leu Val Phe Asn
41 65          70          75          80
42 gct gaa tat gga aac agc tca gtt ttc ttg gag aac agt aca ttt gat      288
43 Ala Glu Tyr Gly Asn Ser Ser Val Phe Leu Glu Asn Ser Thr Phe Asp
44          85          90          95
45 gag ttt gga cat tct atc aat gat tat tca ata tct cct gat ggg cag      336
46 Glu Phe Gly His Ser Ile Asn Asp Tyr Ser Ile Ser Pro Asp Gly Gln
47          100          105          110
48 ttt att ctc tta gaa tac aac tac gtg aag caa tgg agg cat tcc tac      384
49 Phe Ile Leu Leu Glu Tyr Asn Tyr Val Lys Gln Trp Arg His Ser Tyr
50          115          120          125
51 aca gct tca tat gac att tat gat tta aat aaa agg cag ctg att aca      432

```

## RAW SEQUENCE LISTING

DATE: 02/08/2005

PATENT APPLICATION: US/10/522,789

TIME: 11:42:55

Input Set : N:\Rasheed\PTO.SR.txt

Output Set: N:\CRF4\02082005\J522789.raw

52	Thr	Ala	Ser	Tyr	Asp	Ile	Tyr	Asp	Leu	Asn	Lys	Arg	Gln	Leu	Ile	Thr	
53		130					135					140					
54	gaa	gag	agg	att	cca	aac	aac	aca	cag	tgg	gtc	aca	tgg	tca	cca	gtg	480
55	Glu	Glu	Arg	Ile	Pro	Asn	Asn	Thr	Gln	Trp	Val	Thr	Trp	Ser	Pro	Val	
56	145					150					155				160		
57	ggt	cat	aaa	ttg	gca	tat	gtt	tgg	aac	aat	gac	att	tat	gtt	aaa	att	528
58	Gly	His	Lys	Leu	Ala	Tyr	Val	Trp	Asn	Asn	Asp	Ile	Tyr	Val	Lys	Ile	
59				165						170					175		
60	gaa	cca	aat	tta	cca	agt	tac	aga	atc	aca	tgg	acg	ggg	aaa	gaa	gat	576
61	Glu	Pro	Asn	Leu	Pro	Ser	Tyr	Arg	Ile	Thr	Trp	Thr	Gly	Lys	Glu	Asp	
62				180					185					190			
63	ata	ata	tat	aat	gga	ata	act	gac	tgg	gtt	tat	gaa	gag	gaa	gtc	ttc	624
64	Ile	Ile	Tyr	Asn	Gly	Ile	Thr	Asp	Trp	Val	Tyr	Glu	Glu	Glu	Val	Phe	
65		195						200				205					
66	agt	gcc	tac	tct	gct	ctg	tgg	tgg	tct	cca	aac	ggc	act	ttt	tta	gca	672
67	Ser	Ala	Tyr	Ser	Ala	Leu	Trp	Trp	Ser	Pro	Asn	Gly	Thr	Phe	Leu	Ala	
68		210					215				220						
69	tat	gcc	caa	ttt	aac	gac	aca	gaa	gtc	cca	ctt	att	gaa	tac	tcc	ttc	720
70	Tyr	Ala	Gln	Phe	Asn	Asp	Thr	Glu	Val	Pro	Leu	Ile	Glu	Tyr	Ser	Phe	
71	225				230					235					240		
72	tac	tct	gat	gag	tca	ctg	cag	tac	cca	aag	act	gta	cgg	gtt	cca	tat	768
73	Tyr	Ser	Asp	Glu	Ser	Leu	Gln	Tyr	Pro	Lys	Thr	Val	Arg	Val	Pro	Tyr	
74				245					250					255			
75	cca	aag	gca	gga	gct	gtg	aat	cca	act	gta	aag	ttc	ttt	gtt	gta	aat	816
76	Pro	Lys	Ala	Gly	Ala	Val	Asn	Pro	Thr	Val	Lys	Phe	Phe	Val	Val	Asn	
77			260					265				270					
78	aca	gac	tct	ctc	agc	tca	gtc	acc	aat	gca	act	tcc	ata	caa	atc	act	864
79	Thr	Asp	Ser	Leu	Ser	Ser	Val	Thr	Asn	Ala	Thr	Ser	Ile	Gln	Ile	Thr	
80		275					280					285					
81	gct	cct	gct	tct	atg	ttg	ata	ggg	gat	cac	tac	ttg	tgt	gat	gtg	aca	912
82	Ala	Pro	Ala	Ser	Met	Leu	Ile	Gly	Asp	His	Tyr	Leu	Cys	Asp	Val	Thr	
83		290				295				300							
84	tgg	gca	aca	caa	gaa	aga	att	tct	ttg	cag	tgg	ctc	agg	agg	att	cag	960
85	Trp	Ala	Thr	Gln	Glu	Arg	Ile	Ser	Leu	Gln	Trp	Leu	Arg	Arg	Ile	Gln	
86	305				310					315					320		
87	aac	tat	tcg	gtc	atg	gat	att	tgt	gac	tat	gat	gaa	tcc	agt	gga	aga	1008
88	Asn	Tyr	Ser	Val	Met	Asp	Ile	Cys	Asp	Tyr	Asp	Glu	Ser	Ser	Gly	Arg	
89				325					330					335			
90	tgg	aac	tgc	tta	gtg	gca	cgg	caa	cac	att	gaa	atg	agt	act	act	ggc	1056
91	Trp	Asn	Cys	Leu	Val	Ala	Arg	Gln	His	Ile	Glu	Met	Ser	Thr	Thr	Gly	
92			340					345						350			
93	tgg	gtt	gga	aga	ttt	agg	cct	tca	gaa	cct	cat	ttt	acc	ctt	gat	ggt	1104
94	Trp	Val	Gly	Arg	Phe	Arg	Pro	Ser	Glu	Pro	His	Phe	Thr	Leu	Asp	Gly	
95		355					360					365					
96	aat	agc	ttc	tac	aag	atc	atc	agc	aat	gaa	gaa	ggt	tac	aga	cac	att	1152
97	Asn	Ser	Phe	Tyr	Lys	Ile	Ile	Ser	Asn	Glu	Glu	Gly	Tyr	Arg	His	Ile	
98		370				375					380						
99	tgc	tat	ttc	caa	ata	gat	aaa	aaa	gac	tgc	aca	ttt	att	aca	aaa	ggc	1200
100	Cys	Tyr	Phe	Gln	Ile	Asp	Lys	Lys	Asp	Cys	Thr	Phe	Ile	Thr	Lys	Gly	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,789

DATE: 02/08/2005

TIME: 11:42:55

Input Set : N:\Rasheed\PTO.SR.txt

Output Set: N:\CRF4\02082005\J522789.raw

101	385		390		395		400	
102	acc	tgg	gaa	gtc	atc	ggg	ata	gaa
103	Thr	Trp	Glu	Val	Ile	Gly	Ile	Glu
104				405			410	
105	tac	att	agt	aat	gaa	tat	aaa	gga
106	Tyr	Ile	Ser	Asn	Glu	Tyr	Lys	Gly
107				420			425	
108	aaa	atc	caa	ctt	agt	gac	tat	aca
109	Lys	Ile	Gln	Leu	Ser	Asp	Tyr	Thr
110			435				440	
111	ctg	aat	ccg	gaa	agg	tgt	cag	tac
112	Leu	Asn	Pro	Glu	Arg	Cys	Gln	Tyr
113			450				455	
114	gcg	aag	tat	tat	cag	ctg	aga	tgt
115	Ala	Lys	Tyr	Tyr	Gln	Leu	Arg	Cys
116	465						470	
117	act	cta	cac	agc	agc	gtg	aat	gat
118	Thr	Leu	His	Ser	Ser	Val	Asn	Asp
119				485			490	
120	aat	tca	gct	ttg	gat	aaa	atg	ctg
121	Asn	Ser	Ala	Leu	Asp	Lys	Met	Leu
122			500				505	
123	aaa	ctg	gac	ttc	att	att	ttg	aat
124	Lys	Leu	Asp	Phe	Ile	Ile	Leu	Asn
125			515				520	
126	atc	ttg	cct	cct	cat	ttt	gat	aaa
127	Ile	Leu	Pro	Pro	His	Phe	Asp	Lys
128			530				535	
129	gat	gtg	tat	gca	ggc	cca	tgt	agt
130	Asp	Val	Tyr	Ala	Gly	Pro	Cys	Ser
131	545						550	
132	ctg	aac	tgg	gcc	act	tac	ctt	gca
133	Leu	Asn	Trp	Ala	Thr	Tyr	Leu	Ala
134				565			570	
135	agc	ttt	gat	ggc	aga	gga	agt	ggt
136	Ser	Phe	Asp	Gly	Arg	Gly	Ser	Gly
137				580			585	
138	gca	atc	aac	aga	aga	ctg	gga	aca
139	Ala	Ile	Asn	Arg	Arg	Leu	Gly	Thr
140			595				600	
141	gca	gcc	aga	caa	ttt	tca	aaa	atg
142	Ala	Ala	Arg	Gln	Phe	Ser	Lys	Met
143			610				615	
144	gca	att	tgg	ggc	tgg	tca	tat	gga
145	Ala	Ile	Trp	Gly	Trp	Ser	Tyr	Gly
146	625						630	
147	gga	tcg	gga	agt	ggc	gtg	ttc	aag
148	Gly	Ser	Gly	Ser	Gly	Val	Phe	Lys
149				645			650	

## RAW SEQUENCE LISTING

DATE: 02/08/2005

PATENT APPLICATION: US/10/522,789

TIME: 11:42:55

Input Set : N:\Rasheed\PTO.SR.txt

Output Set: N:\CRF4\02082005\J522789.raw

```

150 tcc cgg tgg gag tac tat gac tca gtg tac aca gaa cgt tac atg ggt      2016
151 Ser Arg Trp Glu Tyr Tyr Asp Ser Val Tyr Thr Glu Arg Tyr Met Gly
152          660          665          670
153 ctc cca act cca gaa gac aac ctt gac cat tac aga aat tca aca gtc      2064
154 Leu Pro Thr Pro Glu Asp Asn Leu Asp His Tyr Arg Asn Ser Thr Val
155          675          680          685
156 atg agc aga gct gaa aat ttt aaa caa gtt gag tac ctc ctt att cat      2112
157 Met Ser Arg Ala Glu Asn Phe Lys Gln Val Glu Tyr Leu Leu Ile His
158          690          695          700
159 gga aca gca gat gat aac gtt cac ttt cag cag tca gct cag atc tcc      2160
160 Gly Thr Ala Asp Asp Asn Val His Phe Gln Gln Ser Ala Gln Ile Ser
161 705          710          715          720
162 aaa gcc ctg gtc gat gtt gga gtg gat ttc cag gca atg tgg tat act      2208
163 Lys Ala Leu Val Asp Val Gly Val Asp Phe Gln Ala Met Trp Tyr Thr
164          725          730          735
165 gat gaa gac cat gga ata gct agc agc aca gca cac caa cat ata tat      2256
166 Asp Glu Asp His Gly Ile Ala Ser Ser Thr Ala His Gln His Ile Tyr
167          740          745          750
168 acc cac atg agc cac ttc ata aaa caa tgt ttc tct tta cct tag      2301
169 Thr His Met Ser His Phe Ile Lys Gln Cys Phe Ser Leu Pro
170          755          760          765
173 <210> SEQ ID NO: 2
174 <211> LENGTH: 766
175 <212> TYPE: PRT
176 <213> ORGANISM: Homo sapiens
178 <400> SEQUENCE: 2
180 Met Lys Thr Pro Trp Lys Val Leu Leu Gly Leu Leu Gly Ala Ala Ala
181 1          5          10          15
182 Leu Val Thr Ile Ile Thr Val Pro Val Val Leu Leu Asn Lys Gly Thr
183          20          25          30
184 Asp Asp Ala Thr Ala Asp Ser Arg Lys Thr Tyr Thr Leu Thr Asp Tyr
185          35          40          45
186 Leu Lys Asn Thr Tyr Arg Leu Lys Leu Tyr Ser Leu Arg Trp Ile Ser
187          50          55          60
188 Asp His Glu Tyr Leu Tyr Lys Gln Glu Asn Asn Ile Leu Val Phe Asn
189 65          70          75          80
190 Ala Glu Tyr Gly Asn Ser Ser Val Phe Leu Glu Asn Ser Thr Phe Asp
191          85          90          95
192 Glu Phe Gly His Ser Ile Asn Asp Tyr Ser Ile Ser Pro Asp Gly Gln
193          100          105          110
194 Phe Ile Leu Leu Glu Tyr Asn Tyr Val Lys Gln Trp Arg His Ser Tyr
195          115          120          125
196 Thr Ala Ser Tyr Asp Ile Tyr Asp Leu Asn Lys Arg Gln Leu Ile Thr
197          130          135          140
198 Glu Glu Arg Ile Pro Asn Asn Thr Gln Trp Val Thr Trp Ser Pro Val
199 145          150          155          160
200 Gly His Lys Leu Ala Tyr Val Trp Asn Asn Asp Ile Tyr Val Lys Ile
201          165          170          175
202 Glu Pro Asn Leu Pro Ser Tyr Arg Ile Thr Trp Thr Gly Lys Glu Asp

```

## RAW SEQUENCE LISTING

DATE: 02/08/2005

PATENT APPLICATION: US/10/522,789

TIME: 11:42:55

Input Set : N:\Rasheed\PTO.SR.txt

Output Set: N:\CRF4\02082005\J522789.raw

203		180		185		190
204	Ile	Ile	Tyr	Asn	Gly	Ile
205		195		200		205
206	Ser	Ala	Tyr	Ser	Ala	Leu
207		210		215		220
208	Tyr	Ala	Gln	Phe	Asn	Asp
209	225			230		235
210	Tyr	Ser	Asp	Glu	Ser	Leu
211		245		250		255
212	Pro	Lys	Ala	Gly	Ala	Val
213		260		265		270
214	Thr	Asp	Ser	Leu	Ser	Ser
215		275		280		285
216	Ala	Pro	Ala	Ser	Met	Leu
217		290		295		300
218	Trp	Ala	Thr	Gln	Glu	Arg
219	305			310		315
220	Asn	Tyr	Ser	Val	Met	Asp
221		325		330		335
222	Trp	Asn	Cys	Leu	Val	Ala
223		340		345		350
224	Trp	Val	Gly	Arg	Phe	Arg
225		355		360		365
226	Asn	Ser	Phe	Tyr	Lys	Ile
227		370		375		380
228	Cys	Tyr	Phe	Gln	Ile	Asp
229	385			390		395
230	Thr	Trp	Glu	Val	Ile	Gly
231		405		410		415
232	Tyr	Ile	Ser	Asn	Glu	Tyr
233		420		425		430
234	Lys	Ile	Gln	Leu	Ser	Asp
235		435		440		445
236	Leu	Asn	Pro	Glu	Arg	Cys
237		450		455		460
238	Ala	Lys	Tyr	Tyr	Gln	Leu
239	465			470		475
240	Thr	Leu	His	Ser	Ser	Val
241		485		490		495
242	Asn	Ser	Ala	Leu	Asp	Lys
243		500		505		510
244	Lys	Leu	Asp	Phe	Ile	Ile
245		515		520		525
246	Ile	Leu	Pro	Pro	His	Phe
247		530		535		540
248	Asp	Val	Tyr	Ala	Gly	Pro
249	545			550		555
250	Leu	Asn	Trp	Ala	Thr	Tyr
251		565		570		575

**VERIFICATION SUMMARY**

DATE: 02/08/2005

PATENT APPLICATION: US/10/522,789

TIME: 11:42:56

Input Set : N:\Rasheed\PTO.SR.txt

Output Set: N:\CRF4\02082005\J522789.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:26 M:258 W: Mandatory Feature missing, &lt;223&gt; Blank for SEQ#:1,Line#:24